A/B TESTING SIMULATION ON AIRBNB BARCELONA STAY DURATION

TECHNICAL NOTE

1. Project Objective

This project serves as a practical demonstration of A/B testing, a powerful experimental design methodology widely used in various fields, including product development, marketing, and data science. The primary objective is to showcase the application of A/B testing principles to validate whether a specific variable (or "treatment") significantly impacts a target outcome variable.
In this particular simulation, I explore how a hypothetical feature, such as "flexible cancellation," might influence the average stay duration for Airbnb listings. By designing and executing a controlled experiment, we aim to measure the causal effect of this feature on user behavior.

2. Source of Data

The data used for this simulation is sourced from insideairbnb.com, an independent, non-commercial project that provides open-source datasets for various global cities. These datasets offer detailed information about Airbnb listings, including price, room type, location, number of reviews, and more.
For this project, we specifically utilized the data for *Barcelona, Spain*. The file downloaded was:
listings.csv: This file contains comprehensive details about individual Airbnb accommodations, including attributes such as `id`, `name`, `price`, and `number_of_reviews`, which were crucial for this simulation.

The selection of Barcelona data was made for its rich detail and representativeness, without any specific bias or pre-conceived notions towards its characteristics.

3. Introduction to A/B Testing Theory

A/B testing, also known as split testing or controlled experimentation, is a method of comparing two versions of something (A and B) to determine which one performs better. It is rooted in the principles of statistical hypothesis testing.

Gabriela Inzunza Gómez www.linkedin.com/in/gabriela-inzunza-gomez
g.inzunza.g@seintegra.pro

Core Principles:

- *Randomization*: Participants (in this case, Airbnb listings) are randomly assigned to either Group A (Control) or Group B (Treatment). Randomization ensures that, on average, the two groups are statistically equivalent across all characteristics, both observed and unobserved, before the "treatment" is applied. This minimizes confounding factors and ensures that any observed differences in the outcome can be confidently attributed to the treatment itself.

- *Control Group (A):* This group does not receive the new feature or change. It serves as a baseline against which the performance of the treatment group is measured.

- *Treatment Group (B)*: This group receives the new feature or change being tested.

- *Hypothesis Formulation*:
  - *Null Hypothesis (H0)*: States that there is no statistically significant difference between the control and treatment groups for the outcome metric. Any observed difference is due to random chance.
  - *Alternative Hypothesis (H1)* : States that there is a statistically significant difference between the control and treatment groups, implying that the treatment had an effect.
  - *Statistical Significance*: After collecting data from both groups, statistical tests (like a t-test for comparing means, or a chi-squared test for proportions) are used to calculate a p-value. The p-value indicates the probability of observing the data (or more extreme data) if the null hypothesis were true. If the p-value is below a predetermined significance level (alpha, commonly 0.05), the null hypothesis is rejected, and the difference is deemed statistically significant.

- *Purpose*: A/B testing allows for data-driven decision-making, enabling businesses to validate product changes, marketing campaigns, or feature implementations with statistical confidence before a full rollout.

4. Methodology and Implementation Details

The A/B test simulation was performed using Python, leveraging libraries such as `pandas` for data manipulation, `numpy` for numerical operations and simulations, `scipy.stats` for statistical testing, and `matplotlib` / `seaborn` for data visualization.

The key steps implemented in the accompanying Jupyter Notebook (Airbnb_AB_Test_Barcelona_StayDuration.ipynb) include:

1. *Data Loading & Initial Cleaning:* Importing the `barcelona_listings.csv` dataset, selecting relevant columns, handling missing values, and cleaning the 'price' column to ensure numerical integrity.

2. *Simulated Baseline Stay Duration:* Generating a realistic `simulated_stay` duration for each listing using a Poisson distribution, with the mean influenced by the number of reviews (e.g., listings with more reviews having a slightly longer average stay). This creates a credible baseline for the outcome variable.

3. *A/B Group Assignment:* Randomly assigning each Airbnb listing to either a Control (Group A) or Treatment (Group B) group.

4. *Simulating Treatment Effect:* For the Treatment Group (B), a fixed increase of 1 night was artificially added to their `simulated_stay` duration, representing the hypothetical impact of a feature like "flexible cancellation."

5. *Comparative Analysis:* Calculating and comparing the average stay durations between Group A and Group B.

6. *Statistical Testing:* Performing an independent samples t-test to formally assess if the observed difference in average stay duration between the groups is statistically significant.

7. *Visualization:* Generating histograms to visually compare the distributions of stay duration between the two groups, providing an intuitive understanding of the treatment's effect.

5. Conclusions and Observations

For a detailed walkthrough of the implementation, step-by-step code explanations, numerical results (average stay durations for each group, t-statistic, p-value), and graphical representations, please refer to the accompanying Jupyter Notebook:

Airbnb_AB_Test_Barcelona_StayDuration.ipynb

The notebook demonstrates:

- How to set up an A/B test simulation in a controlled environment.
- The process of calculating and comparing key metrics between control and treatment groups.
- The application of statistical tests to draw conclusions about the significance of observed differences.
- The importance of data visualization in complementing statistical findings and communicating insights.

The results presented in the `.ipynb` file will show whether the simulated "flexible cancellation" feature led to a statistically significant increase in the average stay duration for Group B, as designed in our simulation. This project serves as a robust example of applying data science principles to experimental design.